

PRIVATE AI COMPUTE & SERVICES

Deploy your AI service on dedicated compute.

Dedicated compute to run your applications.
Development and validation support when you need it.

From GPU access to a running AI service.

Need GPU capacity quickly?

AI teams and SIs can agree reserved capacity, an access date and a ready-to-use environment.

Need help delivering the application?

Enterprise teams and software providers can add software, deployment, integration and operations.

Managed Private AI

Bring your application, or ask us to scope software and deployment services.
Combine dedicated compute with the operational support you need.

You / your application partner

Own data, business workflows and acceptance criteria. Agree application responsibilities with us.

RiverSoft

Operate compute. Add software, model deployment, integration and application support as agreed.

Prepare for deployment when needed

Need to develop?

Private AI Workspace

Build and test your application on dedicated compute.

Need to validate?

Private AI Pilot

Evaluate quality, performance and resource use before confirming the deployment configuration.

Ready workloads can move directly to Managed Private AI.

Illustrative deployment: a customer AI assistant

A software provider has built an AI assistant for a customer that needs a dedicated environment. Together, we confirm the model, data access and expected user load. If sizing is uncertain, a Pilot tests the workload. Optional model deployment and integration prepare the service for launch, followed by agreed operations through Managed Private AI.

Each engagement is agreed separately. A Pilot does not commit you to a managed service.

Agree the workload. Define the delivery.

Start with your service requirements and build a proposal around them.

WHAT WE DEFINE TOGETHER	
Workload & goals	Application, models, representative data, expected usage and success criteria.
Compute & environment	GPU resources, CPU, RAM, storage, network and software requirements.
Private access & data	Authorized users, access method, resource isolation, data location and retention.
Operations & support	Responsibilities, service hours, monitoring, response targets, changes and acceptance tests.

Ongoing operation with Managed Private AI

Deploy & accept

Configure compute and access. Validate agreed endpoints and integrations before handover.

Run agreed operations

Use an agreed runbook for monitoring, incident response and maintenance. Review service metrics.

Review capacity

Discuss configuration or capacity changes as your workload and usage evolve.

What you receive

ENGAGEMENT	AGREED DELIVERABLES
Managed Private AI	Environment, access guidance and operating scope. Optional application deployment, integrations and ongoing support as agreed.
Private AI Workspace	Compute access for the agreed term, environment details and basic access guidance.
Private AI Pilot	Results on agreed quality, response time, throughput and resource-use measures, with test conditions and deployment recommendations.

Configuration options. Start with one of our three immediately available reference configurations. Other GPUs, higher-capacity systems and multi-node requirements can be discussed. See our Product Catalog for hardware references.

Commercial scope. Compute, onboarding, software and recurring managed services are priced to the agreed scope. Capacity, delivery date, term and service commitments are confirmed in your proposal.

Document AI: Ask for our existing iDP solution sheet for application details.

Discuss your deployment and request a tailored compute proposal.

www.riversoft.com.tw